

Multi-parameter singular Radon transforms III: real analytic surfaces

Elias M. Stein* and Brian Street†

Abstract

The goal of this paper is to study operators of the form,

$$Tf(x) = \psi(x) \int f(\gamma_t(x)) K(t) dt,$$

where γ is a real analytic function defined on a neighborhood of the origin in $(t, x) \in \mathbb{R}^N \times \mathbb{R}^n$, satisfying $\gamma_0(x) \equiv x$, ψ is a cutoff function supported near $0 \in \mathbb{R}^n$, and K is a “multi-parameter singular kernel” supported near $0 \in \mathbb{R}^N$. A main example is when K is a “product kernel.” We also study maximal operators of the form,

$$\mathcal{M}f(x) = \psi(x) \sup_{0 < \delta_1, \dots, \delta_N < 1} \int_{|t| < 1} |f(\gamma_{\delta_1 t_1, \dots, \delta_N t_N}(x))| dt.$$

We show that \mathcal{M} is bounded on L^p ($1 < p \leq \infty$). We give conditions on γ under which T is bounded on L^p ($1 < p < \infty$); these conditions hold automatically when K is a Calderón-Zygmund kernel. This is the final paper in a three part series. The first two papers consider the more general case when γ is C^∞ .

1 Introduction

In this paper we consider operators of the form,

$$Tf(x) = \psi(x) \int f(\gamma_t(x)) K(t) dt, \quad (1.1)$$

where $\psi \in C_0^\infty(\mathbb{R}^n)$ is supported near 0, $\gamma_t(x) : \mathbb{R}_0^N \times \mathbb{R}_0^n \rightarrow \mathbb{R}^n$ is a germ of a real analytic function (defined on a neighborhood of $(0, 0)$) satisfying $\gamma_0(x) \equiv x$,¹ and K is a “multi-parameter” distribution kernel, supported near $0 \in \mathbb{R}^N$. For instance, one could take K to be a “product kernel” supported near 0.² To define this notion, suppose we have decomposed $\mathbb{R}^N = \mathbb{R}^{N_1} \times \dots \times \mathbb{R}^{N_\nu}$. A product kernel satisfies

$$|\partial_{t_1}^{\alpha_1} \dots \partial_{t_\nu}^{\alpha_\nu} K(t)| \lesssim |t_1|^{-N_1 - |\alpha_1|} \dots |t_\nu|^{-N_\nu - |\alpha_\nu|}, \quad (1.2)$$

along with certain “cancellation conditions” (see Section 16 of [Str11b]).³ We will also study maximal operators of the form,

$$\mathcal{M}f(x) = \sup_{0 < \delta_1, \dots, \delta_\nu \leq a} \psi(x) \int_{|t| \leq 1} |f(\gamma_{(\delta_1 t_1, \dots, \delta_\nu t_\nu)}(x))| dt_1 \dots dt_\nu,$$

*Partially supported by NSF DMS-0901040.

†Partially supported by NSF DMS-0802587.

¹Here we write $f : \mathbb{R}_0^N \rightarrow \mathbb{R}^m$ to denote that f is a germ of a function defined on a neighborhood of 0.

²Our main theorem applies to kernels more general than product kernels.

³The simplest example of a product kernel is given by $K(t_1, \dots, t_\nu) = K_1(t_1) \otimes \dots \otimes K_\nu(t_\nu)$, where K_1, \dots, K_ν are Calderón-Zygmund kernels. That is, they satisfy $|\partial_{t_j}^\alpha K_j(t_j)| \lesssim |t_j|^{-N_j - |\alpha|}$, again along with certain “cancellation conditions.” When $\nu = 1$, the class of product kernels is exactly the class of Calderón-Zygmund kernels. For a precise statement of these cancellation conditions, see Section 16 of [Str11b]; we do not make it precise in this paper, since we deal with more general kernels.

where ψ is as before with $\psi \geq 0$, and $a > 0$ is assumed to be small.

First we describe our results in the single parameter case ($\nu = 1$). In that case, we consider K to be a standard Calderón-Zygmund kernel supported near 0 (a Calderón-Zygmund kernel is the special case of product kernels with $\nu = 1$). In this case, the operator in (1.1) is bounded on L^p ($1 < p < \infty$) with no additional assumptions (provided ψ and K have sufficiently small support, depending on γ). Similarly, when $\nu = 1$, \mathcal{M} is bounded on L^p ($1 < p \leq \infty$) (provided ψ has small enough support and $a > 0$ is sufficiently small).

When we move to the multi-parameter case, the study of T and \mathcal{M} diverge. The results for \mathcal{M} are simple to state: just as in the single-parameter case, \mathcal{M} is bounded on L^p ($1 < p \leq \infty$) with no additional assumptions. In fact, this will follow from proving the L^p boundedness for even stronger maximal operators.

Unfortunately, the results for T are not so simple. Indeed, for $\gamma : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$, given by $\gamma_{(s,t)}(x) = x - st$, there are product kernels $K(s, t)$ of arbitrarily small support such that T is not bounded on L^2 (this was first noted in [NW77], see also Section 17.5 of [Str11b]). Thus, it is necessary to introduce additional assumptions on γ to obtain the L^p boundedness of T .

We now describe a special case of our results, for ν -parameter product kernels $K(t_1, \dots, t_\nu)$ (thus satisfying (1.2)). In [CNSW99] it was shown that γ could be written asymptotically in the form,⁴

$$\gamma_t(x) \sim \exp\left(\sum_{|\alpha| > 0} t^\alpha X_\alpha\right)x, \quad (1.3)$$

where each X_α is a real analytic vector field. Separate each multi-index $\alpha = (\alpha_1, \dots, \alpha_\nu)$, where $\alpha_\mu \in \mathbb{N}^{N_\mu}$ is a multi-index, and $t^\alpha = t_1^{\alpha_1} \cdots t_\nu^{\alpha_\nu}$. We call α a *pure power* if $\alpha_\mu \neq 0$ for only one μ . Otherwise, we call α a *non-pure power*.

A special case of our theorem is as follows: if $X_\alpha = 0$ for every non-pure power α , then T is bounded on L^p ($1 < p < \infty$). In the single parameter case, every power is a pure power, and so this subsumes the single-parameter result. In fact, we will be able to deal with some cases when the non-pure powers are not necessarily zero. Our assumption will be that the pure powers “control” the non-pure powers, in an appropriate sense.

This paper is the third in a series. The first two [Str11b, SS11b] dealt with the more general situation when γ is C^∞ , instead of real analytic. The theorems in those papers took a rather complicated form. We will see that after an appropriate “preparation theorem,”⁵ the main result in this paper for the singular Radon transform T is actually a special case of the results in [SS11b]. The idea is that when γ is assumed to be real analytic, many of the assumptions in [SS11b] come for free. See [SS11a] for an announcement of this series, and an overview tying all three papers together.

The maximal operator \mathcal{M} is not a special case of the results in [SS11b]. Indeed, we will prove a new maximal result concerning $C^\infty \gamma$ (see Section 7) which will imply the maximal result for \mathcal{M} . While this result was not covered in [SS11b], we will see that many of the methods can be transferred to this situation, and the main outline of the proof is quite similar.

2 A motivating special case

In this section we explain our argument in a special motivating case, which contains an essential point which we will use (in various forms) throughout the paper.

First we describe a special case of the results in [CNSW99] in the C^∞ context. Indeed, suppose for each multi-index α , $0 < |\alpha| \leq L$, we are given a C^∞ vector fields X_α defined on a neighborhood of 0. Suppose that this collection of vector fields satisfies Hörmander’s condition at 0: the set of X_α along

⁴(1.3) simply means $\gamma_t(x) = \exp\left(\sum_{0 < |\alpha| < M} t^\alpha X_\alpha\right)x + O(|t|^M)$, for every M as $t \rightarrow 0$. In particular, the reader may wish to consider just the case when $\gamma_t(x) = \exp\left(\sum_{0 < |\alpha| \leq L} t^\alpha X_\alpha\right)x$, and the X_α are germs of real analytic vector fields.

⁵The preparation theorem we need is a Weierstrass type preparation theorem due to Galligo [Gal79].

with all their commutators of all orders spans the tangent space at 0. Define a function γ by,

$$\gamma_t(x) = \exp \left(\sum_{0 < |\alpha| \leq L} t^\alpha X_\alpha \right) x.$$

It is a theorem of Christ, Nagel, Stein, and Wainger [CNSW99] that the operator,

$$f \mapsto \psi(x) \int f(\gamma_t(x)) K(t) dt \quad (2.1)$$

is bounded on L^p ($1 < p < \infty$), for every standard Calderón-Zygmund kernel supported on a sufficiently small neighborhood of 0, and for $\psi \in C_0^\infty$, supported on a sufficiently small neighborhood of 0.

It was discussed in Section 3 of [Str11b] that one need not assume the X_α satisfy Hörmander's condition. Instead, one may assume the weaker condition that the involutive distribution generated by the X_α is locally finitely generated as a C^∞ module. For then, the standard Frobenius theorem holds and foliates the ambient space into leaves; the X_α satisfying Hörmander's condition on each leaf. As was discussed in Section 3 of [Str11b], the methods of [CNSW99] are not sufficient to obtain the L^p boundedness of (2.1) in this case. Nevertheless, the L^p ($1 < p < \infty$) boundedness holds and is a special case of the results in [SS11b].

Now we specialize to the case when the vector fields X_α are real analytic. The involutive distribution generated by a finite collection of real analytic vector fields is *always* locally finitely generated as a C^∞ module. This fact seems to have first been noted in [Nag66, Lob70], see Section 9 for a further discussion. Thus, when the vector fields are real analytic, the L^p boundedness of (2.1) holds. This idea is the core of this entire paper, and similar arguments will be used throughout.

3 Kernels

In this section, we will discuss the classes of kernels $K(t)$ for which we will study operators of the form (1.1). The kernels which we study will be supported in $B^N(a)$, where $a > 0$ is some small number to be chosen later (depending on γ). Fix $\nu \in \mathbb{N}$, we will be studying ν parameter operators.

We suppose we are given ν -parameter dilations on \mathbb{R}^N . That is, we are given $e = (e_1, \dots, e_N)$, with each $0 \neq e_j = (e_j^1, \dots, e_j^\nu) \in \mathbb{N}^\nu$ (here, $0 \in \mathbb{N}$). For $\delta \in [0, \infty)^\nu$ and $t = (t_1, \dots, t_N) \in \mathbb{R}^N$, we define,⁶

$$\delta t = (\delta^{e_1} t_1, \dots, \delta^{e_N} t_N), \quad (3.1)$$

thereby obtaining ν -parameter dilations on \mathbb{R}^N . For each μ , $1 \leq \mu \leq \nu$, let t_μ denote those coordinates t_j of $t = (t_1, \dots, t_N) \in \mathbb{R}^N$ such that $e_j^\mu \neq 0$.

The class of distributions we will define depends on N , a , e , and ν . Given a function ς on \mathbb{R}^N , and $j \in \mathbb{N}^\nu$, define,

$$\varsigma^{(2^j)}(t) = 2^{j \cdot e_1 + \dots + j \cdot e_N} \varsigma(2^j t).$$

Note that $\varsigma^{(2^j)}$ is defined in such a way that,

$$\int \varsigma^{(2^j)}(t) dt = \int \varsigma(t) dt.$$

Definition 3.1. We define $\mathcal{K} = \mathcal{K}(N, e, a, \nu)$ to be the set of all distributions, K , of the form

$$K = \sum_{j \in \mathbb{N}^\nu} \varsigma_j^{(2^j)}, \quad (3.2)$$

where $\{\varsigma_j\}_{j \in \mathbb{N}^\nu} \subset C_0^\infty(B^N(a))$ is a bounded set, satisfying

$$\int \varsigma_j(t) dt_\mu = 0, \quad 0 \neq j_\mu.$$

It was shown in [Str11b] that any sum of the form (3.2) converges in the sense of distributions.

⁶Here δ^{e_j} is defined via standard multi-index notation: $\delta^{e_j} = \prod_\mu \delta_\mu^{e_j^\mu}$.

See [Str11b] for a more in-depth discussion of the class \mathcal{K} .

Remark 3.2. The class of kernels studied in [SS11b] was slightly more general: it was allowed to depend on another parameter μ_0 , $1 \leq \mu_0 \leq \nu$, and the coordinates of each e_j could be elements of $[0, \infty)$, instead of \mathbb{N} . The results in this paper can be extended to deal with that case as well (with essentially no additional work), but we state the results in this simpler case for clarity. See Section 12 for some comments on this.

4 Multi-parameter Carnot-Carathéodory geometry

To state our theorem regarding the singular Radon transforms given by (1.1) in full generality, we must introduce the notion of Carnot-Carathéodory geometry; this notion played an essential role in [Str11b, SS11b]. Our main reference for Carnot-Carathéodory geometry is [Str11a], and we refer the reader there for more information. In this section we introduce only the most basic definitions associated with Carnot-Carathéodory geometry.

Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let X_1, \dots, X_q be C^∞ vector fields on Ω . Denote this list of vector fields by X . We define the Carnot-Carathéodory ball, centered at $x_0 \in \Omega$, of unit radius, with respect to the list of vector fields X by

$$B_X(x_0) := \left\{ y \in \Omega \left| \begin{aligned} &\exists \gamma : [0, 1] \rightarrow \Omega, \gamma(0) = x_0, \gamma(1) = y, \\ &\gamma'(t) = \sum_{j=1}^q a_j(t) X_j(\gamma(t)), a_j \in L^\infty([0, 1]), \\ &\left\| \left(\sum_{1 \leq j \leq q} |a_j|^2 \right)^{\frac{1}{2}} \right\|_{L^\infty([0, 1])} < 1 \end{aligned} \right. \right\}.$$

Now that we have the definition of balls with unit radius, we may define (multi-parameter) balls of any radius merely by scaling the vector fields. To do so, we assign to each vector field, X_j , a (multi-parameter) formal degree $0 \neq d_j = (d_j^1, \dots, d_j^\nu) \in \mathbb{N}^\nu$. For $\delta = (\delta_1, \dots, \delta_\nu) \in [0, \infty)^\nu$, we define the list of vector fields δX to be the list $(\delta^{d_1} X_1, \dots, \delta^{d_q} X_q)$. Here, δ^{d_j} is defined by the standard multi-index notation: $\delta^{d_j} = \prod_{\mu=1}^\nu \delta_\mu^{d_j^\mu}$. We define the ball of radius δ centered at $x_0 \in \Omega$ by

$$B_{(X, d)}(x_0, \delta) := B_{\delta X}(x_0).$$

Definition 4.1. Let $(X, d) = (X_1, d_1), \dots, (X_q, d_q)$ be a finite list of C^∞ vector fields with multi-parameter formal degrees as above. Fix $x_0 \in \Omega$. Let (X_0, d_0) be another C^∞ vector field with multi-parameter formal degree $0 \neq d_0 \in \mathbb{N}^\nu$. We say that (X, d) *controls* (X_0, d_0) on a neighborhood of x_0 if there exists an open set U with $x_0 \in U \subseteq \Omega$, and $\tau_1 > 0$ such that for every $\delta \in [0, 1]^\nu$, $x \in U$, there exist $c_{x, j}^\delta \in C^0(B_{(X, d)}(x, \tau_1 \delta))$ ($1 \leq j \leq q$) such that,

- $\delta^{d_0} X_0 = \sum_{j=1}^q c_{x, j}^\delta \delta^{d_j} X_j$, on $B_{(X, d)}(x, \tau_1 \delta)$.
- $\sup_{\delta \in [0, 1]^\nu} \sum_{|\alpha| \leq m} \|(\delta X)^\alpha c_{x, j}^\delta\|_{C^0(B_{(X, d)}(x, \tau_1 \delta))} < \infty$, for every $m \in \mathbb{N}$.⁷

Note that, since τ_1 and U may be chosen as small as we wish, this is a local property.

Definition 4.2. Let \mathcal{S} be a, possibly infinite, set of germs of C^∞ vector fields, X , defined on a neighborhood of $x_0 \in \mathbb{R}^n$ each paired with a nonzero formal degree $0 \neq d \in \mathbb{N}^\nu$. Let (X_0, d_0) be another germ of a C^∞ vector field defined on a neighborhood of x_0 , with formal degree $0 \neq d_0 \in \mathbb{N}^\nu$. We say \mathcal{S} *controls* (X_0, d_0) on a neighborhood of x_0 if there is a finite subset $\mathcal{F} \subseteq \mathcal{S}$ such that \mathcal{F} controls (X_0, d_0) on a neighborhood of x_0 in the sense of Definition 4.1.

⁷For an arbitrary set $U \subseteq \mathbb{R}^n$, we define $\|f\|_{C^0(U)} = \sup_{y \in U} |f(y)|$, and if we say $\|f\|_{C^0(U)} < \infty$, we mean that f is continuous on U and the norm is finite.

Remark 4.3. Much more detailed information on this notion of control can be found in Section 5.3 of [Str11a] and Section 11.1 of [Str11b].

Definition 4.4. Let \mathcal{S} be a set of germs of C^∞ vector fields defined on a neighborhood of $x_0 \in \mathbb{R}^n$ each paired with a ν -parameter formal degree $0 \neq d \in \mathbb{N}^\nu$. We define $\mathcal{L}(\mathcal{S})$ to be the smallest set of germs of vector fields with formal degrees such that:

- $\mathcal{S} \subseteq \mathcal{L}(\mathcal{S})$,
- if $(X_1, d_1), (X_2, d_2) \in \mathcal{L}(\mathcal{S})$ then $([X_1, X_2], d_1 + d_2) \in \mathcal{L}(\mathcal{S})$.

Furthermore, define $\mathcal{L}_0(\mathcal{S})$ to be the smallest set of germs of vector fields with formal degrees such that:

- $\mathcal{S} \subseteq \mathcal{L}_0(\mathcal{S})$,
- if $(X_1, d_1) \in \mathcal{S}$ and $(X_2, d_2) \in \mathcal{L}_0(\mathcal{S})$, then $([X_1, X_2], d_1 + d_2) \in \mathcal{L}_0(\mathcal{S})$.

Remark 4.5. Note, by the Jacobi identity, for every $(Y_0, d_0) \in \mathcal{L}(\mathcal{S})$,

$$Y_0 \in \text{span}\{Y : (Y, d_0) \in \mathcal{L}_0(\mathcal{S})\}.$$

5 Results

We begin by rigorously stating our maximal result. Let $\gamma(t, x) = \gamma_t(x) : \mathbb{R}_0^N \times \mathbb{R}_0^n \rightarrow \mathbb{R}^n$ be a real analytic function defined on a neighborhood of $(0, 0) \in \mathbb{R}^N \times \mathbb{R}^n$, satisfying $\gamma_0(x) \equiv x$.

For $\psi \in C_0^\infty(\mathbb{R}^n)$ supported on a sufficiently small neighborhood of 0, $\psi \geq 0$, and $a > 0$ sufficiently small, define,

$$\mathcal{M}f(x) = \sup_{0 < \delta_1, \dots, \delta_N \leq 1} \psi(x) \int_{|t| < a} |f(\gamma_{\delta_1 t_1, \dots, \delta_N t_N}(x))| dt.$$

Theorem 5.1. \mathcal{M} is bounded $L^p \rightarrow L^p$ ($1 < p \leq \infty$), provided a is taken sufficiently small, and ψ is supported on a sufficiently small neighborhood of 0.

Theorem 5.1 will follow from a more general maximal theorem about $C^\infty \gamma$ which is discussed in Section 7. This theorem will imply maximal results for even stronger maximal functions than are covered by Theorem 5.1.

Fix ν -parameter dilations $e = (e_1, \dots, e_N)$ on \mathbb{R}^N , as in Section 3 (so that $0 \neq e_j \in \mathbb{N}^\nu$). For a multi-index $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$, define,

$$\deg(\alpha) = \sum_{j=1}^N e_j \alpha_j \in \mathbb{N}^\nu.$$

Definition 5.2. We call α a *pure power* if $\deg(\alpha)$ is nonzero in precisely one component. Otherwise we call α a *non-pure power*.

Let γ be as above. As discussed in the introduction, [CNSW99] showed that γ could be written asymptotically as,

$$\gamma_t(x) \sim \exp\left(\sum_{0 < |\alpha|} t^\alpha X_\alpha\right)x,$$

where the X_α are real analytic vector fields. Define two sets,

$$\begin{aligned} \mathcal{P} &= \{(X_\alpha, \deg(\alpha)) : \alpha \text{ is a pure power}\}, \\ \mathcal{N} &= \{(X_\alpha, \deg(\alpha)) : \alpha \text{ is a non-pure power}\}. \end{aligned} \tag{5.1}$$

Theorem 5.3. *Suppose that for every $(X, d) \in \mathcal{N}$, $\mathcal{L}(\mathcal{P})$ controls (X, d) on a neighborhood of 0.⁸ Then, there exists $a > 0$ such that for every $\psi_1, \psi_2 \in C_0^\infty(\mathbb{R}^n)$ supported on a sufficiently small neighborhood of 0, every $K \in \mathcal{K}(N, e, a, \nu)$, and every C^∞ function $\kappa(t, x)$, the operator given by*

$$Tf(x) = \psi_1(x) \int f(\gamma_t(x)) \psi_2(\gamma_t(x)) \kappa(t, x) K(t) dt \quad (5.2)$$

is bounded $L^p \rightarrow L^p$ ($1 < p < \infty$).

Remark 5.4. Note, by taking ψ_2 to be equal to 1 on a neighborhood of the support of ψ_1 , taking $\kappa = 1$, and taking $a > 0$ so small that for t in the support of $K(t)$ and x in the support of $\psi_1(x)$ we have $\psi_2(\gamma_t(x)) = 1$, we see that the operator given by (1.1) is of the form discussed in Theorem 5.3.

Theorem 5.3 will follow from a more general theorem about C^∞ γ , which is proven in [SS11b].

Corollary 5.5. *Suppose that $X_\alpha = 0$ for every non-pure power α . Then the operator given by (5.2) is bounded on L^p ($1 < p < \infty$).*

Proof. It follows immediately from the definitions that $\mathcal{L}(\mathcal{P})$ controls $(0, \deg(\alpha))$ for every α . Thus the hypotheses of Theorem 5.3 hold trivially. \square

Corollary 5.6. *In the special case $\nu = 1$ (i.e., the single-parameter case, when $K(t)$ is a Calderón-Zygmund kernel), the operator given by (5.2) is bounded on L^p ($1 < p < \infty$).*

Proof. In the single-parameter case, every α is a pure power. Thus, the hypotheses of Corollary 5.5 hold vacuously in this case. \square

Proposition 5.7. *Suppose that T_1 and T_2 are operators of the form covered in Theorem 5.3:*

$$T_j f(x) = \psi_1^j(x) \int f(\gamma_{t_j}^j(x)) \psi_2^j(\gamma_{t_j}^j(x)) \kappa_j(t, x) K_j(t) dt_j, \quad j = 1, 2,$$

with T_j satisfying all of the hypotheses of Theorem 5.3 (with perhaps different dilations e for $\gamma_{t_1}^1$ and $\gamma_{t_2}^2$). Then, $T_1 T_2$ and T_1^ satisfy the hypotheses Theorem 5.3 (provided the K_j and ψ_1^j, ψ_2^j have sufficiently small support).*

Of course, Proposition 5.7 does not lead to any new L^p boundedness results: since T_1 and T_2 are bounded on L^p ($1 < p < \infty$) the same is true for $T_1 T_2$ (similarly for T_1^*). What Proposition 5.7 does tell us is that our assumptions are robust enough that one cannot use our theorem, plus algebraic manipulation of the operators in question, to create new operators which are bounded on L^p , but to which our theorem does not apply. Proposition 5.7 is proved at the end of Section 10.1.

6 Past work

There are a number of papers concerning singular and maximal Radon transforms. We review in this section a few which are closely related to our results. The results mentioned here served as motivation for the results in this paper.

One of the first results which comes to mind when considering maximal Radon transforms associated to real analytic curves is the following result of Bourgain [Bou89]: Theorem 5.1 holds in the special case when $\gamma: \mathbb{R}_0 \times \mathbb{R}_0^2 \rightarrow \mathbb{R}^2$ is given by $\gamma_t(x) = x + tv(x)$ and v is a germ of a real analytic vector field on \mathbb{R}^2 . In a manner completely analogous to this paper, Bourgain proves a more general maximal theorem about C^∞ curves. This more general maximal theorem can be seen as a special case of the maximal result in [SS11b].

The paper which served as the primary motivation for the methods in [Str11b, SS11b] was due to Christ, Nagel, Stein, and Wainger [CNSW99], which discussed the single-parameter case (i.e., when $\nu = 1$). As discussed in [Str11b], the methods in [CNSW99] are not sufficient to obtain Corollary 5.6.

⁸The particular neighborhood used in the definition of control may depend on (X, d) .

Nevertheless, Christ, Nagel, Stein, and Wainger were able to obtain a differentiation theorem for real analytic γ , see Section 21 of [CNSW99]. Namely, for any $f \in L^p$ ($1 < p \leq \infty$) supported sufficiently close to 0,

$$f(x) = \lim_{r \rightarrow 0} c_N^{-1} r^{-N} \int_{|t| \leq r} f(\gamma_t(x)) dt, \quad \text{a.e.}$$

As is well known, this follows from Theorem 5.1. In fact, it follows from the weaker result where one takes the supremum over all $\delta_1 = \delta_2 = \dots = \delta_N$.

In fact, the basic idea of the proof of the differentiation theorem in [CNSW99] is closely related to the results in this paper. Indeed, the result follows by applying the Frobenius theorem to show that the ambient space is foliated into leaves, and other results from [CNSW99] could be applied to each leaf to obtain the differentiation theorem. The main reason why our results are stronger than those in [CNSW99] in the single-parameter case, is that we have access to a stronger form of the Frobenius theorem: the one developed in [Str11a].

The last paper we wish to mention is due to Christ [Chr92]; in it, the “strong maximal function associated to a nilpotent Lie group” is discussed. Let G be a connected, simply connected, nilpotent Lie group. Let X_1, \dots, X_N be left invariant vector fields on G . Define $\gamma : \mathbb{R}^N \times G \rightarrow \mathbb{R}^N$ by

$$\gamma_{t_1, \dots, t_N}(x) = e^{t_1 X_1 + \dots + t_N X_N} x.$$

Note that we may choose coordinates so that γ is real analytic. Define a maximal function by,

$$\widehat{\mathcal{M}}f(x) = \sup_{0 < \delta_1, \dots, \delta_N} \int_{|t| \leq 1} |f(\gamma_{\delta_1 t_1, \dots, \delta_N t_N}(x))| dt.$$

It is shown in [Chr92] that $\widehat{\mathcal{M}}$ is bounded on L^p ($1 < p \leq \infty$). There are a few differences between $\widehat{\mathcal{M}}$ and the maximal operator discussed in Theorem 5.1. First, $\widehat{\mathcal{M}}$ does not involve a cutoff function ψ ; this is due to the translation invariance of $\widehat{\mathcal{M}}$, and is not an essential point. Second, the supremum in the definition of $\widehat{\mathcal{M}}$ is over all $\delta_1, \dots, \delta_N$, while in Theorem 5.1 we restrict attention to $\delta_1, \dots, \delta_N$ small. The reason the results in [Chr92] can be stated for all δ is that they are lifted to a setting where there exist global dilations so that the result for all δ follows from the result for small δ ; and so this is not an essential point either. Thus, the L^p boundedness of $\widehat{\mathcal{M}}$ is essentially a special case of Theorem 5.1.

In fact, [Chr92] studies even stronger maximal functions than $\widehat{\mathcal{M}}$. While these are not a special case of Theorem 5.1, they are a special case of the maximal function discussed in Section 7. Thus, the results in [Chr92] are a special case of the results in this paper. Moreover, the methods in [Chr92] provided the main motivation for the results in Section 7.

7 A more general maximal function

In this section, we introduce a stronger maximal theorem. In Section 10.2, we will show that this maximal theorem implies Theorem 5.1.

Before we introduce this maximal theorem, we must explain the connection between germs of C^∞ functions satisfying $\gamma_0(x) \equiv x$, and certain vector fields. Indeed, given a germ of a C^∞ function $\gamma_t(x)$, defined on a neighborhood of $(0, 0) \in \mathbb{R}^N \times \mathbb{R}^n$, and satisfying $\gamma_0(x) \equiv x$, it makes sense to consider $\gamma_t^{-1}(x)$, since for t sufficiently small, γ_t is a diffeomorphism onto its image.

Thus, we may define,

$$W(t, x) = \frac{d}{d\epsilon} \bigg|_{\epsilon=1} \gamma_{\epsilon t} \circ \gamma_t^{-1}(x) \in T_x \mathbb{R}^n.$$

Note that $W(t)$ is vector field, depending smoothly on t such that $W(0) \equiv 0$.

Proposition 7.1 (Proposition 12.1 of [Str11b]). *The map $\gamma \mapsto W$ is a bijection between germs of C^∞ functions, as above, to germs of vector fields, $W(t)$, depending smoothly on t and satisfying $W(0) \equiv 0$.*

Proof sketch. The inverse of the map $\gamma \mapsto W$ is as follows. Given W , let $\omega(\epsilon, t, x)$ be the unique solution to the ODE:

$$\frac{d}{d\epsilon} \omega(\epsilon, t, x) = \frac{1}{\epsilon} W(\epsilon t, \omega(\epsilon, t, x)), \quad \omega(0, t, x) = x.$$

Define $\gamma_t(x) = \omega(1, t, x)$.⁹ This map $W \mapsto \gamma$ is the two-sided inverse to the map $\gamma \mapsto W$. See Proposition 12.1 of [Str11b] for details. \square

In light of Proposition 7.1, instead of defining γ , we may instead define W . This will allow us to introduce dilations on γ_t that are not of the form $\gamma(\delta_1 t_1, \dots, \delta_N t_N)$, thereby allowing us to introduce stronger maximal functions than are covered in Theorem 5.1. A similar idea was used in [Chr92], though the setting was simpler and the vector field W did not need to be introduced.

We now turn to defining the maximal function. Let $(X, d) = (X_1, d_1), \dots, (X_q, d_q)$ be germs of C^∞ vector fields defined on a neighborhood of $0 \in \mathbb{R}^n$, each with an associated formal degree $0 \neq d_j \in \mathbb{N}^\nu$. We suppose that $([X_j, X_k], d_j + d_k)$ is controlled by (X, d) on a neighborhood of 0, for every $1 \leq j, k \leq q$. Let $1 \leq r \leq q$, and suppose each d_j , $1 \leq j \leq r$ is nonzero in only one component. Suppose further that $(X_1, d_1), \dots, (X_r, d_r)$ generate $(X_{r+1}, d_{r+1}), \dots, (X_q, d_q)$ in the sense that $(X_{r+1}, d_{r+1}), \dots, (X_q, d_q) \in \mathcal{L}_0(\{(X_1, d_1), \dots, (X_r, d_r)\})$.

We suppose we are given ν -parameter dilations $e = (e_1, \dots, e_N)$ on \mathbb{R}^N , as in Section 3; thus it makes sense to write δt for $t \in \mathbb{R}^N$ and $\delta \in [0, 1]^\nu$ (see (3.1)).

We suppose we are given germs of C^∞ functions,

$$c_j(t, s, x) : \mathbb{R}_0^N \times \mathbb{R}_0^N \times \mathbb{R}_0^n \rightarrow \mathbb{R}, \quad j = 1, \dots, q,$$

with $c_j(0, 0, x) \equiv 0$. Suppose $0 \neq \alpha_1, \dots, \alpha_r \in \mathbb{N}^\nu$ be multi-indices such that,

$$\frac{1}{\alpha_l!} \frac{\partial^{\alpha_l}}{\partial s} c_l(t, s, x) \Big|_{t=s=0} = 1, \quad 1 \leq l \leq r, \quad (7.1)$$

and for all $1 \leq l \leq r$, $1 \leq k \leq q$ and all β_1, β_2 with $\beta_1 + \beta_2 = \alpha_l$,

$$\frac{\partial^{\beta_1}}{\partial t} \frac{\partial^{\beta_2}}{\partial s} c_k(t, s, x) \Big|_{t=s=0} = 0, \quad \text{unless } l = k, \beta_1 = 0, \beta_2 = \alpha_l. \quad (7.2)$$

Let $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$. For each $j \in \mathbb{N}_\infty^\nu$, define,

$$W_j(t, x) = \sum_{l=1}^q c_l(2^{-j}t, t, x) 2^{-j \cdot d_l} X_l.$$

Given W_j we obtain a corresponding γ_t^j as in Proposition 7.1. That is, let ω_j be the unique solution to the ODE:

$$\frac{d}{d\epsilon} \omega_j(\epsilon, t, x) = \frac{1}{\epsilon} W_j(\epsilon t, \omega_j(\epsilon, t, x)), \quad \omega_j(0, t, x) = x.$$

Set $\gamma_t^j(x) = \omega_j(1, t, x)$. It is easy to see, via the contraction mapping principle, that there are open sets $0 \in U \subset \mathbb{R}^N$, $0 \in V \subset \mathbb{R}^n$, independent of j , such that $\gamma_t^j : U \times V \rightarrow \mathbb{R}^n$.

Let $\psi_1, \psi_2 \in C_0^\infty(\mathbb{R}^n)$ be supported on a small neighborhood of 0, $\psi_1, \psi_2 \geq 0$, and let $a > 0$ be a small number. In light of the above remarks, it makes sense to define the maximal function,

$$\widetilde{\mathcal{M}}f(x) = \sup_{j \in \mathbb{N}^\nu} \psi_1(x) \int_{|t| < a} \left| f(\gamma_t^j(x)) \right| \psi_2(\gamma_t^j(x)) dt.$$

Theorem 7.2. *Under the above conditions $\widetilde{\mathcal{M}}$ is bounded on L^p ($1 < p \leq \infty$), provided ψ_1 and ψ_2 are supported on a sufficiently small neighborhood of 0, and $a > 0$ is sufficiently small.*

Theorem 7.2 is proved in Section 11.

⁹It is easy to see, via the contraction mapping principle, that the solution ω exists up to $\epsilon = 1$ for t sufficiently small.

Remark 7.3. We will see that Theorem 5.1 follows from Theorem 7.2. It is not hard to see that Theorem 2.4 of [Chr92] follows from Theorem 7.2. It follows that all of the results of [Chr92] can be reduced to Theorem 7.2.

Remark 7.4. Notice that we have discretized our maximal functions; i.e., we only consider dyadic scales. This is essential when considering $\widetilde{\mathcal{M}}$. Indeed, the obvious non-discretized version need not be bounded on all L^p , $p > 1$. This was noted on the top of page 5 of [Chr92].

8 When γ is C^∞ : the results of [SS11b]

In this section, we review the results of [SS11b]. We will see that Theorem 5.3 is, in fact, a special case of Theorem 5.2 of [SS11b], which we review below. In addition, we rephrase the assumptions of [SS11b] in a few different ways, which will be useful in what follows.

The setting is as follows. We are given a C^∞ function $\gamma_t(x) = \gamma(t, x) : \mathbb{R}_0^N \times \mathbb{R}_0^n \rightarrow \mathbb{R}^n$ satisfying $\gamma_0(x) \equiv x$. The goal is to give conditions on γ such that the operator given by¹⁰

$$Tf(x) = \psi(x) \int f(\gamma_t(x)) K(t) dt, \quad (8.1)$$

is bounded on L^p ($1 < p < \infty$) for every $K \in \mathcal{K}(N, e, a, \nu)$ where a is sufficiently small and ψ is supported on a sufficiently small neighborhood of 0. We think of the ν -parameter dilations e as fixed so that it makes sense to write δt for $\delta \in [0, \infty)^\nu$ and $t \in \mathbb{R}^N$ as in (3.1).

Definition 8.1. Let $(X, d) = (X_1, d_1), \dots, (X_q, d_q)$ be a finite list of C^∞ vector fields with ν -parameter formal degrees $0 \neq d_j \in [0, \infty)^\nu$ as in Section 4. Let $W(t, x) \in T_x \mathbb{R}^n$ be a smooth vector field (defined on a neighborhood of $(0, 0) \in \mathbb{R}^N \times \mathbb{R}^n$), depending smoothly on $t \in \mathbb{R}_0^N$. We say that (X, d) *controls* W on a neighborhood of 0 if there exists an open neighborhood U of $0 \in \mathbb{R}^n$, $\tau_1 > 0$, and $\rho_1 > 0$, such that for every $\delta \in [0, 1]^\nu$, $x_0 \in U$, there exist functions $c_l^{x_0, \delta}$ on $B^N(\rho_1) \times B_{(X, d)}(x_0, \tau_1 \delta)$ satisfying

- $W(\delta t, x) = \sum_{l=1}^q c_l^{x_0, \delta}(t, x) \delta^{d_l} X_l(x)$ on $B^N(\rho_0) \times B_{(X, d)}(x_0, \tau_1 \delta)$.
- $\sup_{\substack{x_0 \in U \\ \delta \in [0, 1]^\nu}} \sum_{|\alpha| + |\beta| \leq m} \left\| (\delta X)^\alpha \partial_t^\beta c_l^{x_0, \delta} \right\|_{C^0(B^N(\rho_1) \times B_{(X, d)}(x_0, \tau_1 \delta))} < \infty$, for every m .

If, instead, \mathcal{S} is an infinite collection of vector fields, then we say \mathcal{S} *controls* W on a neighborhood of 0 if there is a finite subset which controls W .

Remark 8.2. Definition 8.1 is closely related to Definition 4.1. Indeed, note that if (X, d) controls W on a neighborhood of 0, and if the Taylor series for W is given by,

$$W(t, x) \sim \sum_{\alpha} t^\alpha Y_\alpha,$$

then (X, d) controls $(Y_\alpha, \deg(\alpha))$ on a neighborhood of 0 for every α .

Definition 8.3. Given a finite list of C^∞ vector fields (defined on a neighborhood of $0 \in \mathbb{R}^n$) with ν -parameter formal degrees $(X_1, d_1), \dots, (X_r, d_r)$ we say that this list *generates the finite list* $(X, d) = (X_1, d_1), \dots, (X_q, d_q)$ (here $q \geq r$) if there exist vector fields with formal degrees $(X_{r+1}, d_{r+1}), \dots, (X_q, d_q) \in \mathcal{L}_0(\{(X_1, d_1), \dots, (X_r, d_r)\})$ such that for every $1 \leq j, k \leq q$, $([X_j, X_k], d_j + d_k)$ is controlled by (X, d) on a neighborhood of 0.

Remark 8.4. In what follows, we will say A *controls* B to mean A controls B on a neighborhood of 0.

With the above definitions in hand, we are prepared to state the assumptions placed on γ in [SS11b]. We state these assumptions in three different ways, which we will see are all equivalent. Under any of

¹⁰Or more generally, operators of the form covered in Theorem 5.3.

the following assumptions, the operator given by (8.1) is bounded on L^p ($1 < p < \infty$). In what follows, define the vector field,

$$W(t, x) = \frac{d}{d\epsilon} \Big|_{\epsilon=1} \gamma_{\epsilon t} \circ \gamma_t^{-1}(x) \in T_x \mathbb{R}^n.$$

Note that γ_t^{-1} makes sense, since for t sufficiently small, γ_t is a diffeomorphism onto its image (because $\gamma_0(x) \equiv x$).

(I) Expand W as a Taylor series in the t variable,

$$W(t) \sim \sum_{|\alpha| > 0} t^\alpha \hat{X}_\alpha,$$

where the \hat{X}_α are C^∞ vector fields. We assume that there is a finite subset,

$$\mathcal{F} \subseteq \left\{ \left(\hat{X}_\alpha, \deg(\alpha) \right) : \alpha \text{ is a pure power} \right\},$$

such that \mathcal{F} generates a finite list (\hat{X}, d) and (\hat{X}, d) controls W .

(II) For the second equivalent condition, we rephrase (I) as having two distinct parts:

(II.F) A “finite type” condition: taking \hat{X}_α as in (I), we assume that there is a finite subset,

$$\mathcal{F} \subseteq \left\{ \left(\hat{X}_\alpha, \deg(\alpha) \right) : \alpha \in \mathbb{N}^N \right\},$$

such that \mathcal{F} generates a finite list (\hat{X}, d) and this finite list controls W .

(II.A) An “algebraic” condition: we assume that for every non-pure power α , $(\hat{X}_\alpha, \deg(\alpha))$ is controlled by

$$\mathcal{L} \left(\left\{ \left(\hat{X}_\alpha, \deg(\alpha) \right) : \alpha \text{ is a pure power} \right\} \right).$$

(III) The third equivalent condition is the same as (II), except we use different vector fields. Indeed, write,

$$\gamma_t(x) \sim \exp \left(\sum_{\alpha} t^\alpha X_\alpha \right) x.$$

(III.F) A “finite type” condition: we assume there is a finite subset,

$$\mathcal{F} \subseteq \{ (X_\alpha, \deg(\alpha)) : \alpha \in \mathbb{N}^\nu \},$$

such that \mathcal{F} generates a finite list (X, d) and this finite list controls W .

(III.A) An “algebraic” condition: we assume that for every non-pure power α , $(X_\alpha, \deg(\alpha))$ is controlled by

$$\mathcal{L}(\{ (X_\alpha, \deg(\alpha)) : \alpha \text{ is a pure power} \}).$$

Remark 8.5. Note that the assumptions of Theorem 5.3 are exactly that (III.A) holds.

Remark 8.6. The vector fields X_α and \hat{X}_α are closely related. See Lemma 8.15.

Theorem 8.7. $(I) \Leftrightarrow (II) \Leftrightarrow (III)$; i.e., the above three conditions are equivalent.

We prove Theorem 8.7 at the end of this section.

Theorem 8.8 (Theorem 5.2 of [SS11b]). *Under any of the above three conditions, there exists a $\delta > 0$ such that for every $\psi_1, \psi_2 \in C_0^\infty(\mathbb{R}^n)$ supported on a sufficiently small neighborhood of 0, every $K \in \mathcal{K}(N, e, a, \nu)$, and every C^∞ function $\kappa(t, x)$, the operator given by,*

$$Tf(x) = \psi_1(x) \int f(\gamma_t(x)) \psi_2(\gamma_t(x)) \kappa(t, x) K(t) dt$$

is bounded $L^p \rightarrow L^p$ ($1 < p < \infty$).

Proof. Under the assumption (I), this is contained in Theorem 5.2 of [SS11b]. \square

Proposition 8.9. *When γ is real analytic, (II.F) and (III.F) hold automatically.*

We defer the proof of Proposition 8.9 to Section 10. From the above results, Theorem 5.3 follows easily.

Proof of Theorem 5.3 given the above results. By Theorem 8.8, it suffices to show that (under the assumptions of Theorem 5.3), (III) holds. Proposition 8.9 shows that (III.F) holds, while the assumptions of Theorem 5.3 are exactly that (III.A) holds. \square

We close this section by proving Theorem 8.7. We separate Theorem 8.7 into two propositions.

Proposition 8.10. *(I) \Leftrightarrow (II).*

Proposition 8.11. *(II) \Leftrightarrow (III). More specifically, (II.F) \Leftrightarrow (III.F) and (II.A) \Leftrightarrow (III.A).*

Lemma 8.12. *The notion of control is transitive. Indeed, if \mathcal{S}_1 and \mathcal{S}_2 are sets of vector fields with ν -parameter formal degrees such that every element of \mathcal{S}_2 is controlled by \mathcal{S}_1 , and if \mathcal{S}_2 controls a vector field with formal degree (X, d) , then so does \mathcal{S}_1 . A similar result holds if (X, d) is replaced by $W(t, x)$ as in Definition 8.1.*

Proof. This follows immediately from the definitions. \square

Lemma 8.13. *If $\mathcal{S}_1, \mathcal{S}_2$ are sets of vector fields with ν -parameter formal degrees, such that,*

- *for every $(X_1, d_1), (X_2, d_2) \in \mathcal{S}_1$, $([X_1, X_2], d_1 + d_2)$ is controlled by \mathcal{S}_1 ,*
- *every element of \mathcal{S}_2 is controlled by \mathcal{S}_1 .*

Then, every element of $\mathcal{L}(\mathcal{S}_2)$ is controlled by \mathcal{S}_1 .

Proof. This follows immediately from the definitions. \square

Remark 8.14. Note that if \mathcal{F} is a finite set of vector fields which generates a finite list $(X, d) = (X_1, d_1), \dots, (X_q, d_q)$ as in Definition 8.3, then (X, d) satisfies the hypotheses of \mathcal{S}_1 in Lemma 8.13.

Proof of Proposition 8.10. (I) \Rightarrow (II): (II.F) follows immediately from (I). (II.A) follows from (I) via Remark 8.2.

(II) \Rightarrow (I): Take \mathcal{F} as in (II) and let \mathcal{F}' be a finite list generated by \mathcal{F} (see Definition 8.3), so that \mathcal{F}' controls W , and set

$$\widehat{\mathcal{P}} = \left\{ \left(\widehat{X}_\alpha, \deg(\alpha) : \alpha \text{ is a pure power} \right) \right\}.$$

By our assumption, every element of \mathcal{F} is controlled by $\mathcal{L}(\widehat{\mathcal{P}})$. Thus, every element of $\mathcal{L}(\mathcal{F})$ is controlled by $\mathcal{L}(\widehat{\mathcal{P}})$ (Lemma 8.13). It follows that every element of \mathcal{F}' is controlled by $\mathcal{L}(\widehat{\mathcal{P}})$. By Remark 4.5, every element of \mathcal{F}' is therefore controlled by $\mathcal{L}_0(\widehat{\mathcal{P}})$. Let $\mathcal{F}_0 \subseteq \mathcal{L}_0(\widehat{\mathcal{P}})$ be a finite subset such that every element of \mathcal{F}' is controlled by \mathcal{F}_0 . We may assume that $\mathcal{F}_0 \subseteq \mathcal{L}_0(\mathcal{F}_0 \cap \widehat{\mathcal{P}})$; indeed, since $\mathcal{F}_0 \subseteq \mathcal{L}_0(\widehat{\mathcal{P}})$, we may add a finite number of elements to \mathcal{F}_0 from $\widehat{\mathcal{P}}$ so that $\mathcal{F}_0 \subseteq \mathcal{L}_0(\mathcal{F}_0 \cap \widehat{\mathcal{P}})$. Since \mathcal{F}' controls W on a neighborhood of 0, it follows that \mathcal{F}_0 controls W (Lemma 8.12).

To complete the proof, we need to show that if $(Y_1, d_1), (Y_2, d_2) \in \mathcal{F}_0$, then $([Y_1, Y_2], d_1 + d_2)$ is controlled by \mathcal{F}_0 ; for then $\mathcal{F}_0 \cap \widehat{\mathcal{P}}$ will generate the finite list \mathcal{F}_0 (which we know to control W). To do this, it suffices to show that $([Y_1, Y_2], d_1 + d_2)$ is controlled by \mathcal{F}' (by Lemma 8.12, since every element of \mathcal{F}' is controlled by \mathcal{F}_0). In particular, it suffices to show that every element of $\mathcal{L}(\widehat{\mathcal{P}})$ is controlled by \mathcal{F}' .

By Remark 8.2, every element of $\widehat{\mathcal{P}}$ is controlled by \mathcal{F}' . We know, by assumption, that if $(X_1, d_1), (X_2, d_2) \in \mathcal{F}'$, then $([X_1, X_2], d_1 + d_2)$ is controlled by \mathcal{F}' . It follows from Lemma 8.13 that every element of $\mathcal{L}(\widehat{\mathcal{P}})$ is controlled by \mathcal{F}' . This completes the proof. \square

Lemma 8.15. *Let \widehat{X}_α be as in (I) and X_α be as in (III), then, for every $d_0 \in \mathbb{N}^\nu$,*

$$\text{span} \{Y : (Y, d_0) \in \mathcal{L}(\{(X_\alpha, \deg(\alpha))\})\} = \text{span} \left\{ Y : (Y, d_0) \in \mathcal{L} \left(\left\{ \left(\widehat{X}_\alpha, \deg(\alpha) \right) \right\} \right) \right\}$$

Proof. This follows easily from an application of the Campbell-Hausdorff formula. See the proof of Proposition 9.6 of [CNSW99] for a similar result and more details. \square

Proof of Proposition 8.11. We begin by showing (II.F) \Rightarrow (III.F); the implication (III.F) \Rightarrow (II.F) follows in the same way, and we leave the details to the reader. Suppose that (II.F) holds: there is a finite set \mathcal{F} as in (II.F) which generates a finite list, \mathcal{F}' and this finite list controls W . Define,

$$\mathcal{S} = \{(X_\alpha, \deg(\alpha)) : \alpha \in \mathbb{N}\}.$$

By Lemma 8.15, for every $(Y_0, d_0) \in \mathcal{F}'$, $Y_0 \in \text{span} \{Y : (Y, d_0) \in \mathcal{L}(\mathcal{S})\}$. Thus, there is a finite subset $\mathcal{F}_0 \subseteq \mathcal{L}(\mathcal{S})$, such that \mathcal{F}_0 controls \mathcal{F}' . By Remark 4.5, we may assume that $\mathcal{F}_0 \subseteq \mathcal{L}_0(\mathcal{S})$. Furthermore, by adding a finite number of elements to \mathcal{F}_0 , we may assume that $\mathcal{F}_0 \subseteq \mathcal{L}_0(\mathcal{S} \cap \mathcal{F}_0)$.

Since \mathcal{F}' controls W , we have that \mathcal{F}_0 controls W . To complete the proof, we need only verify that for every $(X_1, d_1), (X_2, d_2) \in \mathcal{F}_0$, $([X_1, X_2], d_1 + d_2)$ is controlled by \mathcal{F}_0 ; for then $\mathcal{F}_0 \cap \mathcal{S}$ will generate the finite list \mathcal{F}_0 which we already know controls W . That this is true follows just as in the proof of Proposition 8.10. This completes the proof of (III.F).

(II.A) \Leftrightarrow (III.A) is a simple consequence of Lemma 8.15, which we leave to the reader. \square

9 A primer on real analytic functions

In this section, we introduce the theory we need to see our theorems concerning real analytic γ as special cases of theorems concerning C^∞ γ , which are amenable to the methods of [SS11b]. Let

$$\mathcal{A}_N = \{f : \mathbb{R}_0^N \rightarrow \mathbb{R} \mid f \text{ is real analytic}\},$$

the set of germs of real analytic functions defined on a neighborhood of $0 \in \mathbb{R}^N$, and taking values in \mathbb{R} . Note that \mathcal{A}_N is a ring, and,

$$\mathcal{A}_N^m = \{f : \mathbb{R}_0^N \rightarrow \mathbb{R}^m \mid f \text{ is real analytic}\}.$$

\mathcal{A}_N^m is an \mathcal{A}_N -module.

Theorem 9.1. *Suppose $f(t, x) : \mathbb{R}_0^N \times \mathbb{R}_0^n \rightarrow \mathbb{R}^m$ is a germ of a real analytic function; $f \in \mathcal{A}_{N+n}^m$. For $\alpha \in \mathbb{N}^\nu$, let $f_\alpha(x) \in \mathcal{A}_n^m$ be the Taylor coefficient of f , when the Taylor series is taken in the t variable:*

$$f(t, x) = \sum_{\alpha \in \mathbb{N}^\nu} t^\alpha f_\alpha(x). \quad (9.1)$$

Then, there exist finitely many multi-indices $\alpha_1, \dots, \alpha_r \in \mathbb{N}^\nu$, and germs of real analytic functions $c_{\alpha_1}, \dots, c_{\alpha_r} \in \mathcal{A}_{N+n}$ such that,

$$f(t, x) = \sum_{k=1}^r c_{\alpha_k}(t, x) t^{\alpha_k} f_{\alpha_k}(x), \quad (9.2)$$

on a neighborhood of $(0, 0) \in \mathbb{R}^N \times \mathbb{R}^n$. Furthermore, we may assume for every $1 \leq j, k \leq r$,

$$\left. \frac{1}{\alpha_j!} \frac{\partial^{\alpha_j}}{\partial t} t^{\alpha_k} c_{\alpha_k}(t, x) \right|_{t=0} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \quad (9.3)$$

Theorem 9.2. *Suppose*

$$\mathcal{S} \subseteq \mathcal{A}_N^n \times \mathbb{N}^\nu.$$

Then there exists a finite subset $\mathcal{F} \subseteq \mathcal{S}$ such that every $(g, e) \in \mathcal{S}$ can be written in the form,

$$g(x) = \sum_{\substack{(f,d) \in \mathcal{F} \\ d \leq e}} c_{(f,d)}(x) f(x); \quad (9.4)$$

where $c_{(f,d)} \in \mathcal{A}_N$, and $d \leq e$ means that the inequality holds for each coordinate. The neighborhood on which (9.4) holds may depend on (g, e) .

Corollary 9.3. *Let $\mathcal{S} \subseteq \mathcal{A}_N^n \times \mathbb{N}^\nu$. We think of \mathcal{S} as a set of pairs (X, d) where X is the germ of a real analytic vector field, and $d \in \mathbb{N}^\nu$ is a formal degree. Then, there exists a finite subset $\mathcal{F} \subseteq \mathcal{S}$ such that every $(Y, e) \in \mathcal{S}$ is controlled by \mathcal{F} .*

Proof. Let \mathcal{F} be as in the conclusion of Theorem 9.2 when applied to \mathcal{S} . Let $(Y, e) \in \mathcal{S}$. We wish to show that (Y, e) is controlled by \mathcal{F} . For $\delta \in [0, 1]^\nu$, multiplying both sides of (9.4) by δ^e , we obtain,

$$\delta^e Y = \sum_{\substack{(X,d) \in \mathcal{F} \\ d \leq e}} (\delta^{e-d} c_{(X,d)}) \delta^d X.$$

Noting that $\delta^{e-d} c_{(X,d)} \in C^\infty$ uniformly for $\delta \in [0, 1]^\nu$ (since we are only considering the case when $d \leq e$ coordinatewise), the result follows. \square

The above three results are the only results we will need concerning real analytic functions. The rest of this section is devoted to proving and discussing Theorems 9.1 and 9.2. The reader uninterested in the proofs may safely skip the remainder of this section, as it will not be used in the sequel.

Theorems 9.1 and 9.2 both follow easily from well-known results. We begin by outlining the results necessary to prove Theorem 9.2.

Proposition 9.4 (See [ZS75]). *The ring \mathcal{A}_N is Noetherian.*

Comments on the proof. This is a simple consequence of the Weierstrass preparation theorem. See page 148 of [ZS75]. The proof in [ZS75] is for the formal power series ring, however, as mentioned on page 130 of [ZS75], the proof also works for the ring convergent power series: i.e., the ring of power series with some positive radius of convergence. The ring of germs of real analytic functions is isomorphic to the ring convergent power series. \square

Proposition 9.5. *The module \mathcal{A}_N^m is a Noetherian \mathcal{A}_N -module.*

Comments on the proof. It is easy to see that for any Noetherian ring R , the R -module R^m is Noetherian. Actually, in this special case, one can characterize a finite set of generators for any submodule of \mathcal{A}_N^m . This can be found in [Gal79], but we will not need this. \square

Proof of Theorem 9.2. Let $\mathcal{S} \subseteq \mathcal{A}_N^n \times \mathbb{N}^\nu$. Define a map $\iota : \mathcal{A}_N^n \times \mathbb{N}^\nu \rightarrow \mathcal{A}_{\nu+N}^n$ by,

$$\iota(f, d) = t^d f(x), \quad t \in \mathbb{R}^\nu.$$

Let M be the submodule of $\mathcal{A}_{\nu+N}^n$ generated by $\iota\mathcal{S}$. M is finitely generated by Proposition 9.5. Let $\mathcal{F} \subseteq \mathcal{S}$ be a finite subset such that $\iota\mathcal{F}$ generates M . We will show that \mathcal{F} satisfies the conclusions of the theorem.

Indeed, let $(g, e) \in \mathcal{S}$. Since $t^e g(x) \in M$, we may write,

$$t^e g(x) = \sum_{(f,d) \in \mathcal{F}} \widehat{c}_{(f,d)}(t, x) t^d f(x), \quad (9.5)$$

on a neighborhood of $(0, 0) \in \mathbb{R}^\nu \times \mathbb{R}^N$.

We apply $\frac{1}{e!} \frac{\partial^e}{\partial t} \Big|_{t=0}$ to both sides of (9.5). Note that,

$$\frac{1}{e!} \frac{\partial^e}{\partial t} \Big|_{t=0} \widehat{c}_{(f,d)}(t, x) t^d = 0, \quad \text{unless } e \geq d.$$

Thus, we obtain,

$$g(x) = \sum_{\substack{(f,d) \in \mathcal{F} \\ d \leq e}} \left[\frac{1}{e!} \frac{\partial^e}{\partial t} \Big|_{t=0} \widehat{c}_{(f,d)}(t, x) t^d \right] f(x),$$

completing the proof. \square

Remark 9.6. The case $\nu = 0$ of Corollary 9.3 in the context of vector fields seems to have been first used by Lobry [Lob70]. In [Lob70], Corollary 9.3 was used in the following way. Let \mathcal{S} be a set of germs of real analytic vector fields, and let \mathcal{D} be the involutive distribution generated by \mathcal{S} . In light of the $\nu = 0$ case of Corollary 9.3, there is a finite subset $\mathcal{F} \subseteq \mathcal{D}$ such that for every $Y \in \mathcal{D}$, Y can be written as a sum of elements of \mathcal{F} (on some suitably small open set, depending on Y). Because of this, [Lob70] said the distribution \mathcal{D} was “locally of finite type.” Unfortunately, there is a slight error in the application the Frobenius theorem in [Lob70], see [Ste80]—this is due to the fact that the open set depends on Y . However, in our uses of the Frobenius theorem, we will always be able to consider only *finite* sets \mathcal{S} , and then it is easy to see that the open set need not depend on Y . Corollary 9.3 can be considered a “scale invariant” version of the ideas of [Lob70].

We close this section by proving Theorem 9.1. We will need a Weierstrass-type preparation theorem from [Gal79]. First, we introduce the relevant aspects of [Gal79] we need (which is only a small fraction of that paper), and then we will show that Theorem 9.1 is a simple consequence. We will only need part of Theorem 1.2.5 of [Gal79], and we turn to introducing the relevant notation. We will introduce a division theorem for functions in \mathcal{A}_ν^m (of course division theorems are closely related to preparation theorems). Pick numbers $\lambda_1, \dots, \lambda_\nu \in (0, \infty)$ such that $\lambda_1, \dots, \lambda_\nu$ are linearly independent over \mathbb{Z} . For $\alpha = (\alpha_1, \dots, \alpha_\nu) \in \mathbb{N}^\nu$, define $L(\alpha) = \sum_{j=1}^\nu \alpha_j \lambda_j \in [0, \infty)$. Note that if $\alpha \neq \beta$, then $L(\alpha) \neq L(\beta)$. L induces a total ordering on the set $\mathbb{N}^\nu \times \{1, \dots, m\}$. Indeed, we say $(\alpha, i) < (\beta, j)$ if $L(\alpha) < L(\beta)$ or if $L(\alpha) = L(\beta)$ and $i < j$.

For a function $f \in \mathcal{A}_\nu^m$ write $f = (f_1, \dots, f_m)$, where $f_i \in \mathcal{A}_\nu$. Write f_i as a Taylor series,

$$f_i(x) = \sum_{\alpha \in \mathbb{N}^\nu} f_{\alpha,i} x^\alpha.$$

Let $Q(f)$ be the Newton diagram of f :

$$Q(f) = \{(\alpha, i) \in \mathbb{N}^\nu \times \{1, \dots, m\} : f_{\alpha,i} \neq 0\}.$$

For $f \neq 0$, let $\exp_L(f)$ to be the smallest element of $Q(f)$ in the above defined total ordering. Let M be a submodule of \mathcal{A}_ν^m . We define,

$$E_L(M) = \{\exp_L(f) : 0 \neq f \in M\}.$$

Theorem 9.7 (Part of Theorem 1.2.5 of [Gal79]). *Let M be a submodule of \mathcal{A}_ν^m . Then, every element $f \in \mathcal{A}_\nu^m$ is congruent, modulo M , to a unique element $r = r(f) = (r_1, \dots, r_m) \in \mathcal{A}_\nu^m$ of the form,*

$$r_i(x) = \sum_{(\alpha,i) \notin E_L(M)} h_{\alpha,i} x^\alpha. \quad (9.6)$$

That is, the nonzero terms in the Taylor series of r do not appear in $E_L(M)$.

Proof of Theorem 9.1. First we prove the result without insisting the c_{α_k} satisfy (9.3). Then we will show that we may modify the c_{α_k} so that (9.3) is satisfied.

Express f as a Taylor series as in (9.1):

$$f(t, x) = \sum_{\alpha \in \mathbb{N}^N} t^\alpha f_\alpha(x).$$

Let M be the submodule of \mathcal{A}_{N+n}^m generated by $\{t^\alpha f_\alpha(x) : \alpha \in \mathbb{N}^N\}$. We know that M is finitely generated by Proposition 9.5, and thus (9.2) will follow if we can show $f \in M$.

Taking the setup of Theorem 9.7 (and thus we must choose some L), we see that we may write f uniquely modulo M as a term r satisfying (9.6). We wish to show that $r = 0$. Suppose not. We will show that $\exp_L(r) \in E_L(M)$, which will contradict the form of r given by (9.6).

Note that $r = m + f$ for some $m \in M$. We claim that there exists $K > 0$ sufficiently large such that,

$$\exp_L(m + f) = \exp_L\left(m + \sum_{|\alpha| \leq K} t^\alpha f_\alpha(x)\right).$$

Indeed, if $|\alpha|$ is so large that¹¹ $L(\alpha, 0) > L(\exp_L(m + f))$, then the term $t^\alpha f_\alpha(x)$ does not affect $\exp_L(m + f)$.

Note, though, that $m + \sum_{|\alpha| \leq K} t^\alpha f_\alpha(x) \in M$. Thus, by definition, $\exp_L(r) = \exp_L(m + f) \in E_L(M)$. This achieves the contradiction and completes the proof of (9.2).

Now we turn to showing that the c_{α_k} may be modified so that they satisfy (9.3). Indeed, suppose c_{α_k} satisfy (9.2). Define \hat{c}_{α_k} by,

$$t^{\alpha_k} \hat{c}_{\alpha_k}(t, x) = t^{\alpha_k} c_{\alpha_k}(t, x) - \sum_{j=1}^r \frac{t^{\alpha_j}}{\alpha_j!} \left[\frac{\partial^{\alpha_j}}{\partial s} \Big|_{s=0} s^{\alpha_k} c_{\alpha_k}(s, x) \right] + t^{\alpha_k};$$

note that the right hand side is clearly of the form $t^{\alpha_k} \hat{c}_{\alpha_k}$ for some \hat{c}_{α_k} , since $\frac{\partial^{\alpha_j}}{\partial s} \Big|_{s=0} s^{\alpha_k} \hat{c}_{\alpha_k}(s, x) = 0$ unless $\alpha_k \leq \alpha_j$ coordinatewise.

It is clear that \hat{c}_{α_k} satisfies (9.3). Thus, to complete the proof, we need only show that,

$$\sum_{j=1}^r \hat{c}_{\alpha_j}(t, x) t^{\alpha_j} f_{\alpha_j}(x) = f(t, x). \quad (9.7)$$

Since (9.7) holds with \hat{c}_{α_j} replaced by c_{α_j} , it suffices to show,

$$\sum_{k=1}^r t^{\alpha_k} f_{\alpha_k}(x) = \sum_{k=1}^r \left(\sum_{j=1}^r \frac{t^{\alpha_j}}{\alpha_j!} \left[\frac{\partial^{\alpha_j}}{\partial s} \Big|_{s=0} s^{\alpha_k} c_{\alpha_k}(s, x) \right] \right) f_{\alpha_k}(x). \quad (9.8)$$

In light of (9.2), both sides of (9.8) are equal to,

$$\sum_{j=1}^r \frac{t^{\alpha_j}}{\alpha_j!} \frac{\partial^{\alpha_j}}{\partial s} \Big|_{s=0} f(s, x).$$

This completes the proof. \square

10 Reduction to the C^∞ case

In this section, we use the results from Section 9 to reduce Theorems 5.3 and 5.1 to theorems about C^∞ γ : namely, Theorems 8.8 and 7.2. In Section 10.1, we reduce Theorem 5.3 to Theorem 8.8 (or, more precisely, Proposition 8.9), while in Section 10.2 we reduce Theorem 5.1 to Theorem 7.2.

¹¹Here, we are thinking of $(\alpha, 0) \in \mathbb{N}^N \times \mathbb{N}^n$. Also, when we write $L(\exp_L(m + f))$, we are dropping off the last coordinate of $\exp_L(m + f)$ so that the expression makes sense.

10.1 Singular Radon Transforms

In this section, we will show that Theorem 5.1 follows from Theorem 8.8. In fact, as shown in Section 8, it suffices to prove Proposition 8.9: that (II.F) and (III.F) hold automatically when γ is real analytic. Furthermore, since Proposition 8.11 shows that (II.F) and (III.F) are equivalent, it suffices to show that (II.F) holds whenever γ is real analytic.

Lemma 10.1. *Let $\mathcal{F} = \{(X_1, d_1), \dots, (X_r, d_r)\}$ be a finite set of germs of real analytic vector fields (defined on a neighborhood of 0), each paired with formal degree $0 \neq d_j \in \mathbb{N}^\nu$. Then \mathcal{F} generates a finite list, as in Definition 8.3: that is, there is a finite set of elements $(X_{r+1}, d_{r+1}), \dots, (X_q, d_q) \in \mathcal{L}_0(\mathcal{F})$ such that for every $1 \leq i, j \leq q$, $([X_i, X_j], d_i + d_j)$ is controlled by $(X_1, d_1), \dots, (X_q, d_q)$.*

Proof. Apply Corollary 9.3 to $\mathcal{L}(\mathcal{F})$ to obtain a finite set $\mathcal{F}_0 \subseteq \mathcal{L}(\mathcal{F})$ such that \mathcal{F}_0 controls every element of $\mathcal{L}(\mathcal{F})$. By Remark 4.5, we may assume $\mathcal{F}_0 \subseteq \mathcal{L}_0(\mathcal{F})$. We may, without loss of generality, replace \mathcal{F}_0 with $\mathcal{F}_0 \cup \mathcal{F}$. We claim \mathcal{F}_0 is the desired set $(X_1, d_1), \dots, (X_q, d_q)$. Indeed, it only remains to show that for every $(Y_1, f_1), (Y_2, f_2) \in \mathcal{F}_0$, $([Y_1, Y_2], f_1 + f_2)$ is controlled by \mathcal{F}_0 . Since $([Y_1, Y_2], f_1 + f_2) \in \mathcal{L}(\mathcal{F})$, this follows by the definition of \mathcal{F}_0 , completing the proof. \square

Proof of Proposition 8.9. We take W as in Section 8. Since we are assuming γ is real analytic, W is real analytic. We express W as a Taylor series in the t variable:

$$W(t) = \sum_{|\alpha| > 0} t^\alpha \hat{X}_\alpha,$$

where the \hat{X}_α are real analytic vector fields. The goal is to show that there is a finite set,

$$\mathcal{F} \subseteq \left\{ \left(\hat{X}_\alpha, \deg(\alpha) \right) \right\}, \quad (10.1)$$

such that \mathcal{F} generates a finite list, and this finite list controls W . Since the vector fields are real analytic, Lemma 10.1 shows that \mathcal{F} automatically generates a finite list. Thus, it suffices to show that there is a finite set \mathcal{F} as in (10.1) such that \mathcal{F} controls W . We apply Theorem 9.1 to W to show that there exist $\alpha_1, \dots, \alpha_r$ such that,

$$W(t, x) = \sum_{j=1}^r c_j(t, x) t^{\alpha_j} X_{\alpha_j}(x).$$

From here, it is immediate to verify that $(X_{\alpha_1}, \deg(\alpha_1)), \dots, (X_{\alpha_r}, \deg(\alpha_r))$ control W , completing the proof. \square

We close this section by proving Proposition 5.7. The main point is the following. Suppose we are given ν_1 parameter dilations on \mathbb{R}^{N_1} and ν_2 parameter dilations on \mathbb{R}^{N_2} , and suppose we are given $\gamma_{t_j}^j(x) : \mathbb{R}_0^{N_j} \times \mathbb{R}_0^n \rightarrow \mathbb{R}^n$, $j = 1, 2$, germs of real analytic functions, satisfying the hypotheses of Theorem 5.3 (i.e., satisfying (III.A)). Proposition 5.7 will follow if we show that $\gamma_{t_1}^1 \circ \gamma_{t_2}^2(x)$ and $(\gamma_{t_1}^1)^{-1}(x)$ both satisfy the (III.A). For $\gamma_{t_1}^1 \circ \gamma_{t_2}^2$ we are using the $\nu_1 + \nu_2$ parameter dilations on $\mathbb{R}^{N_1+N_2}$ given by, for $(\delta_1, \delta_2) \in [0, 1]^{\nu_1} \times [0, 1]^{\nu_2}$ and $(t_1, t_2) \in \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$, $(\delta_1, \delta_2)(t_1, t_2) = (\delta_1 t_1, \delta_2 t_2)$. We write,

$$\gamma_{t_j}^j(x) \sim \exp \left(\sum_{|\alpha| > 0} t_j^\alpha X_\alpha^j \right) x.$$

Lemma 9.3 of [CNSW99] shows that,

$$(\gamma_{t_1}^1)^{-1}(x) \sim \exp \left(\sum_{|\alpha| > 0} -t_1^\alpha X_\alpha^1 \right) x.$$

The fact that $(\gamma_{t_1}^1)^{-1}$ satisfies (III.A) now follows immediately.

We now turn to $\gamma_{t_1}^1 \circ \gamma_{t_2}^2$. Define

$$\mathcal{T} = \mathcal{L} \left(\left\{ (X_{\alpha_1}^1, (\alpha_1, 0)), (X_{\alpha_2}^2, (0, \alpha_2)) \right\} \right).$$

It follows from the Campbell-Hausdorff formula (see Section 3 of [CNSW99]) that,

$$\gamma_{t_1}^1 \circ \gamma_{t_2}^2 (x) \sim \exp \left(\left[\sum_{|\alpha_1| > 0} t_1^{\alpha_1} X_{\alpha_1}^1 \right] + \left[\sum_{|\alpha_2| > 0} t_2^{\alpha_2} X_{\alpha_2}^2 \right] + \left[\sum_{|\beta_1|, |\beta_2| > 0} t_1^{\beta_1} t_2^{\beta_2} X_{\beta_1, \beta_2} \right] \right) x, \quad (10.2)$$

where $X_{\beta_1, \beta_2} \in \text{span} \{X' : (X', (\beta_1, \beta_2)) \in \mathcal{T}\}$.

Define,

$$\begin{aligned} \mathcal{P}_1 &= \{ (X_{\alpha_1}^1, (\deg(\alpha_1), 0)) : \deg(\alpha_1) \text{ is nonzero in only one component} \} \subseteq \mathcal{A}_n^n \times \mathbb{N}^{\nu_1 + \nu_2}, \\ \mathcal{P}_2 &= \{ (X_{\alpha_2}^2, (0, \deg(\alpha_2))) : \deg(\alpha_2) \text{ is nonzero in only one component} \} \subseteq \mathcal{A}_n^n \times \mathbb{N}^{\nu_1 + \nu_2}; \end{aligned}$$

where $\deg(\alpha_1)$ is defined with the dilations on \mathbb{R}^{N_1} and $\deg(\alpha_2)$ is defined with the dilations on \mathbb{R}^{N_2} . In light of (10.2) the vector fields associated to the pure powers of $\gamma_{t_1}^1 \circ \gamma_{t_2}^2$ are given by $\mathcal{P}_1 \cup \mathcal{P}_2$. Our assumption that γ^1 and γ^2 satisfy (III.A) imply that $(X_{\alpha_1}^1, (\deg(\alpha_1), 0))$ and $(X_{\alpha_2}^2, (0, \deg(\alpha_2)))$ are controlled by $\mathcal{L}(\mathcal{P}_1 \cup \mathcal{P}_2)$ for every α_1 and α_2 . Since every element of \mathcal{T} is given by iterated commutators of $X_{\alpha_1}^1$ and $X_{\alpha_2}^2$, it follows that for every $(X, (\beta_1, \beta_2)) \in \mathcal{T}$, $\mathcal{L}(\mathcal{P}_1 \cup \mathcal{P}_2)$ controls $(X, (\deg(\beta_1), \deg(\beta_2)))$. Hence, $\mathcal{L}(\mathcal{P}_1 \cup \mathcal{P}_2)$ controls $(X_{\beta_1, \beta_2}, (\deg(\beta_1), \deg(\beta_2)))$, for every β_1 and β_2 . Thus, $\gamma_{t_1}^1 \circ \gamma_{t_2}^2$ satisfies (III.A). This completes the proof of Proposition 5.7.

10.2 Maximal Radon transforms

In this section, we reduce Theorem 5.1 to Theorem 7.2. The main tool will be Theorem 9.1.

Let $\gamma : \mathbb{R}_0^N \times \mathbb{R}_0^n \rightarrow \mathbb{R}^n$ be a germ of a real analytic function satisfying $\gamma_0(x) \equiv x$. For $\delta = (\delta_1, \dots, \delta_N) \in [0, 1]^N$ and $t = (t_1, \dots, t_N) \in \mathbb{R}^N$, we define $\delta t = (\delta_1 t_1, \dots, \delta_N t_N)$. The goal is to study the maximal operator,

$$\mathcal{M}f(x) = \sup_{\delta \in [0, 1]^N} \psi_1(x) \int_{|t| \leq a} |f(\gamma_{\delta t}(x))| dt.$$

Where $\psi_1 \in C_0^\infty(\mathbb{R}^n)$ is supported on a small neighborhood of 0 and $\psi_1 \geq 0$. Let $\psi_2 \in C_0^\infty(\mathbb{R}^n)$, $\psi_2 \geq 0$, with $\psi_2 \equiv 1$ on a neighborhood of the support of ψ_1 . We may assume ψ_2 has small support, by shrinking the support of ψ_1 . By taking $a > 0$ small, we may ensure for $|t| < a$, and x in the support of ψ_1 , we have $\psi_2(\gamma_t(x)) = 1$. With this setup, define,

$$\mathcal{M}_0 f(x) = \sup_{j \in \mathbb{N}^N} \psi_1(x) \int_{|t| \leq a} |f(\gamma_{2^{-j}t}(x))| \psi_2(\gamma_{2^{-j}t}(x)) dt.$$

It is easy to see that we have the pointwise inequality, $\mathcal{M}f(x) \lesssim \mathcal{M}_0 f(x)$. Thus, to prove Theorem 5.1, it suffices to prove \mathcal{M}_0 is bounded on L^p , $1 < p \leq \infty$.

Define the real analytic vector field,

$$W(t, x) = \frac{d}{d\epsilon} \Big|_{\epsilon=1} \gamma_{\epsilon t} \circ \gamma_t^{-1}(x) \in T_x \mathbb{R}^n.$$

Note that, for $j \in \mathbb{N}^N$,

$$W(2^{-j}t, x) = \frac{d}{d\epsilon} \Big|_{\epsilon=1} \gamma_{\epsilon 2^{-j}t} \circ \gamma_{2^{-j}t}^{-1}(x) \in T_x \mathbb{R}^n. \quad (10.3)$$

That is, replacing γ_t with $\gamma_{2^{-j}t}$ changes $W(t, x)$ to $W(2^{-j}t, x)$.

Write,

$$W(t, x) = \sum_{|\alpha| > 0} t^\alpha X_\alpha.$$

Applying Theorem 9.1 to $W(t, x)$ we see that there exist $\alpha_1, \dots, \alpha_r$ and germs of real analytic functions c_{α_l} such that,

$$W(t, x) = \sum_{l=1}^r c_{\alpha_l}(t, x) t^{\alpha_l} X_{\alpha_l}. \quad (10.4)$$

Moreover, we may assume that the c_{α_l} satisfy (9.3).

Let $\nu = N + r$. We will define ν -parameter dilations on W . For $l = 1, \dots, r$, let $\widehat{d}_l \in \mathbb{N}^r$ be equal to 1 in the l component and 0 in all other components. Then, for $(j_1, j_2) \in \mathbb{N}_\infty^N \times \mathbb{N}_\infty^r$, define,

$$W_{(j_1, j_2)}(t, x) = \sum_{l=1}^r c_{\alpha_l}(2^{-j_1}t, x) t^{\alpha_l} 2^{-j_2 \cdot \widehat{d}_l} X_{\alpha_l}.$$

Let $\gamma_t^{(j_1, j_2)}$ be the function corresponding to $W_{(j_1, j_2)}$ as in Proposition 7.1. Just as in Section 7, it is easy to see (via the contraction mapping principle) that there exist open sets $0 \in U \subseteq \mathbb{R}^N$, $0 \in V \subseteq \mathbb{R}^n$, independent of j_1, j_2 such that $\gamma^{j_1, j_2} : U \times V \rightarrow \mathbb{R}^n$. By possibly shrinking a and the support of ψ_1, ψ_2 , we may define the maximal function,

$$\mathcal{M}_1 f(x) = \sup_{(j_1, j_2) \in \mathbb{N}^N \times \mathbb{N}^r} \psi_1(x) \int_{|t| \leq a} \left| f\left(\gamma_t^{(j_1, j_2)}(x)\right) \right| \psi_2\left(\gamma_t^{(j_1, j_2)}(x)\right) dt.$$

We claim that $\mathcal{M}_0 f(x) \leq \mathcal{M}_1 f(x)$. To see this, we need only show for every $j \in \mathbb{N}^N$, $\gamma_{2^{-j}t}$ is of the form $\gamma_t^{(j_1, j_2)}$ for some j_1, j_2 . In light of (10.3), it suffices to show for every $j \in \mathbb{N}^N$, $W(2^{-j}t, x)$ is of the form $W_{(j_1, j_2)}(t, x)$ for some j_1, j_2 . In light of (10.4),

$$W(2^{-j}t, x) = \sum_{l=1}^r c_{\alpha_l}(2^{-j}t, x) t^{\alpha_l} 2^{-j \cdot \alpha_l} X_{\alpha_l}.$$

Thus, if we take $j_2 = (j \cdot \alpha_1, j \cdot \alpha_2, \dots, j \cdot \alpha_r)$, we have $W(2^{-j}t, x) = W_{(j, j_2)}(t, x)$. This completes the proof that $\mathcal{M}_0 f(x) \leq \mathcal{M}_1 f(x)$.

Hence, to prove Theorem 5.1, we need only show that \mathcal{M}_1 is bounded on L^p , $1 < p \leq \infty$. We will show that \mathcal{M}_1 is of the form covered in Theorem 7.2, thereby reducing Theorem 5.1 to Theorem 7.2.

For $l = 1, \dots, r$, let $X_l = X_{\alpha_l}$, $c_l(t, s, x) = c_{\alpha_l}(t, x) s^{\alpha_l}$, and $d_l \in \mathbb{N}^\nu = \mathbb{N}^N \times \mathbb{N}^r$ be given by $d_l = (0, \widehat{d}_l) \in \mathbb{N}^N \times \mathbb{N}^r$. Furthermore, for $(j_1, j_2) \in \mathbb{N}^N \times \mathbb{N}^r$, we define $2^{-(j_1, j_2)}t = 2^{-j_1}t$. With this new notation, for $j \in \mathbb{N}^\nu$, we have

$$W_j(t, x) = \sum_{l=1}^r c_l(2^{-j}t, t, x) 2^{-j \cdot d_l} X_l.$$

We apply Lemma 10.1 to extend the list $(X_1, d_1), \dots, (X_r, d_r)$ to a list $(X_1, d_1), \dots, (X_q, d_q)$ as in Lemma 10.1: that this extended list satisfies the hypotheses of the list of the same name in Section 7 is exactly the conclusion of Lemma 10.1.

For $r+1 \leq l \leq q$, define $c_l(t, s, x) \equiv 0$. Note, we have,

$$W_j(t, x) = \sum_{l=1}^q c_l(2^{-j}t, t, x) 2^{-j \cdot d_l} X_l.$$

To complete the proof that \mathcal{M}_1 satisfies the hypotheses of Theorem 7.2, we need only show that c_1, \dots, c_r satisfy (7.1) and c_1, \dots, c_q satisfy (7.2). (7.2) is trivial for c_{r+1}, \dots, c_q (since they are all 0) and so we need only verify (7.1) and (7.2) for c_1, \dots, c_r . Here, we are taking $\alpha_1, \dots, \alpha_r$ as above (see (10.4)). (7.1) and (7.2) will follow from the fact that $c_{\alpha_1}, \dots, c_{\alpha_r}$ satisfy (9.3).

First we verify (7.1). Let $1 \leq l \leq r$. Consider,

$$\frac{1}{\alpha_l!} \frac{\partial^{\alpha_l}}{\partial s} \bigg|_{s=t=0} c_l(t, s, x) = \frac{1}{\alpha_l!} \frac{\partial^{\alpha_l}}{\partial s} \bigg|_{s=t=0} c_{\alpha_l}(t, x) s^{\alpha_l} = \frac{1}{\alpha_l!} \frac{\partial^{\alpha_l}}{\partial t} \bigg|_{t=0} c_{\alpha_l}(t, x) t^{\alpha_l} = 1,$$

where the last equality follows from (9.3). Thus, (7.1) holds.

We turn to (7.2). Fix $1 \leq l, k \leq r$ and β_1, β_2 such that $\beta_1 + \beta_2 = \alpha_l$. Consider,

$$\left. \frac{\partial^{\beta_1}}{\partial t} \frac{\partial^{\beta_2}}{\partial s} \right|_{s=t=0} c_k(t, s, x) = \left. \frac{\partial^{\beta_1}}{\partial t} \frac{\partial^{\beta_2}}{\partial s} \right|_{s=t=0} c_{\alpha_k}(t, x) s^{\alpha_k}. \quad (10.5)$$

Note that the right hand side of (10.5) is 0 unless $\beta_2 = \alpha_k$. Thus, we need only consider the case when $\beta_2 = \alpha_k$; in this case, we have,

$$\left. \frac{\partial^{\beta_1}}{\partial t} \frac{\partial^{\beta_2}}{\partial s} \right|_{s=t=0} c_{\alpha_k}(t, x) s^{\alpha_k} = C \left. \frac{\partial^{\alpha_l}}{\partial t} \right|_{t=0} c_{\alpha_k}(t, x) t^{\alpha_k}, \quad (10.6)$$

where C is some constant. Note that the right hand side of (10.6) is 0 unless $l = k$, by (9.3). (7.2) follows. This completes the proof that \mathcal{M}_1 is of the form covered by Theorem 7.2, and finishes the reduction of Theorem 5.1 to Theorem 7.2.

Remark 10.2. Let us take a moment to remark on the essential idea of this section. When one is considering the singular Radon transform (Theorem 5.3), which vector fields correspond to pure powers and non-pure powers is forced, due to the nature of the cancellation in the singular kernel. However, when we consider the maximal function, we introduce the cancellation in an *ad hoc* way (see the operators B_j in Section 11). Because of this, we have some freedom in choosing which vector fields correspond to pure powers, by considering a stronger maximal operator. This idea was adapted from [Chr92].

11 Proof of the general maximal result (Theorem 7.2)

In this section, we prove Theorem 7.2. The proof is a modification of the proof of Theorem 5.4 of [SS11b]. First we will introduce some necessary auxiliary operators, in a manner completely analogous to the methods in [SS11b]. Then, we will describe the modifications of the proof in [SS11b] necessary to prove Theorem 7.2. The reader may wish to have a copy of [SS11b] at hand, as we will be referring to it repeatedly.

The proof of Theorem 7.2 proceeds by induction on ν . We begin by describing the necessary modifications to Section 9 of [SS11b], where the induction is set up. We take all the same notation as Theorem 7.2. Let $\psi_0 \in C_0^\infty(\mathbb{R}^n)$ be non-negative and satisfy $\psi_1, \psi_2 \prec \psi_0$. We also assume that ψ_0 has small support. Let $\sigma \in C_0^\infty(B^N(a))$ satisfy $\sigma \geq 0$ and $\sigma \geq 1$ on a neighborhood of 0. We define for $j \in \mathbb{N}_\infty^\nu$,

$$M_j f(x) = \psi_0(x) \int f(\gamma_t^j(x)) \psi_0(\gamma_t^j(x)) \sigma(t) dt.$$

It is immediate to see, if we shrink $a > 0$ in the definition of $\widetilde{\mathcal{M}}$, we have,

$$\widetilde{\mathcal{M}} f(x) \lesssim \sup_{j \in \mathbb{N}} M_j |f|(x).$$

Thus, to prove Theorem 7.2 it suffices to prove the following proposition,

Proposition 11.1.

$$\left\| \sup_{j \in \mathbb{N}^\nu} |M_j f(x)| \right\|_{L^p} \lesssim \|f\|_{L^p},$$

for $1 < p < \infty$.

Indeed, merely apply Proposition 11.1 to $|f|$ to prove Theorem 7.2. It is Proposition 11.1 which we prove by induction on ν . For $E \subseteq \{1, \dots, \nu\}$ and $j = (j^1, \dots, j^\nu) \in \mathbb{N}^\nu$, define $j_E = (j_E^1, \dots, j_E^\nu) \in \mathbb{N}_\infty^\nu$ by,

$$j_E^\mu = \begin{cases} j^\mu & \text{if } \mu \in E, \\ \infty & \text{if } \mu \notin E. \end{cases}$$

Thinking of j_E as an element of $\mathbb{N}^{|E|}$ (by suppressing those coordinates which equal ∞), it is easy to see that M_{j_E} is of the same form as M_j , but with ν replaced by E . In particular, $\gamma_t^{j_E}$ is of the same form as γ_t^j , but instead with $|E|$ parameter dilations. Thus our inductive hypothesis implies for $E \subsetneq \{1, \dots, \nu\}$,

$$\left\| \sup_{j \in \mathbb{N}^\nu} |M_{j_E} f(x)| \right\|_{L^p} \lesssim \|f\|_{L^p}, \quad 1 < p < \infty.$$

Note that the base case of our induction is trivial. Indeed, $M_{j_\emptyset} f = \left[\int \sigma(t) dt \right] \psi_0^2 f$.

For $j \in \mathbb{N}_\infty^\nu$, we define A_j from the list of vector fields (X, d) just as in [SS11b]. Similarly, for $j \in \mathbb{N}^\nu$, we define D_j just as in [SS11b]. For $j \in \mathbb{N}^\nu$, we define,

$$B_j = \sum_{E \subseteq \{1, \dots, \nu\}} (-1)^{|E|} A_{j_E^c} M_{j_E}.$$

Just as in [SS11b], Proposition 11.1 follows from the following proposition,

Proposition 11.2.

$$\left\| \sup_{j \in \mathbb{N}^\nu} |B_j f(x)| \right\|_{L^p} \lesssim \|f\|_{L^p},$$

for $1 < p < \infty$.

It follows in exactly the same manner as [SS11b] that to prove Proposition 11.2, it suffices to prove,

Proposition 11.3. *If $a > 0$ is sufficiently small, there exists $\epsilon > 0$ such that,*

$$\|B_j D_k\|_{L^2 \rightarrow L^2} \lesssim 2^{-\epsilon|j-k|},$$

for $j, k \in \mathbb{N}^\nu$.

The proof in [SS11b] of the result analogous to Proposition 11.3 (Theorem 10.1 of [SS11b]) follows by reducing the question to a general result in [Str11b]. The proof of Proposition 11.3 has the same basic outline as the proof of Theorem 10.1 of [SS11b], with only a few minor differences. We outline, below, the necessary facts needed to adapt the proof in [SS11b] to our situation. Note that A_j and D_j are defined in the same way as in [SS11b]—we therefore only need to discuss the modifications necessary to deal with the new form of M_j .

One key point is the following; for $1 \leq l \leq r$, and $j' \in \mathbb{N}_\infty^\nu$,

$$\left. \frac{1}{\alpha_l!} \frac{\partial}{\partial t} \right|_{t=0} W_{j'}(t, x) = 2^{-j' \cdot d_l} X_l; \quad (11.1)$$

i.e., $2^{-j' \cdot d_l} X_l$ is the Taylor coefficient of t^{α_l} , when the Taylor series of $W_{j'}$ is taken in the t variable. This is an immediate consequence of the definition of W_j and (7.1) and (7.2). (11.1) is the main property needed when M_{j_E} plays the role of some S_l in [SS11b].

Remark 11.4. One also needs that “ M_{j_E} is controlled by $(2^{-j \wedge k} X, \sum d)$ at the unit scale” (see [SS11b] for this terminology). This follows immediately from the definition of M_{j_E} .

The other main property we need is as follows. In the case when $j^{\mu_1} - k^{\mu_1} = |j - k|_\infty$, for some μ_1 , we must use $M_{j_E \cup \{\mu_1\}} - M_{j_E}$ as $R_1 - R_2$ in the argument (see [SS11b] for a discussion of what we mean by $R_1 - R_2$). Define, $\widehat{W}_{j,k,E,\mu_1}(t, s, x)$ by the same formula as $W_{j_E}(t, x)$ except with 2^{-j_E} replaced by $\delta = (\delta_1, \dots, \delta_\nu) \in [0, 1]^\nu$, where,

$$\delta_\mu = \begin{cases} 2^{-j_E^\mu} & \text{if } \mu \neq \mu_1, \\ s 2^{-k^{\mu_1}} & \text{if } \mu = \mu_1. \end{cases}$$

Note that,

$$\widehat{W}_{j,k,E,\mu_1}(t, 0, x) = W_{j_E}(t, x), \quad \widehat{W}_{j,k,E,\mu_1}\left(t, 2^{k^{\mu_1} - j^{\mu_1}}, x\right) = W_{j_E \cup \{\mu_1\}}(t, x).$$

Thus, if we let $\widehat{\gamma}_{t,s}(x)$ be the function associated to $\widehat{W}_{j,k,E,\mu_1}(t,s,x)$ as in Proposition 7.1, we see that,

$$M_{j_E} f(x) = \psi_0(x) \int f(\widehat{\gamma}_{t,0}(x)) \psi_0(\widehat{\gamma}_{t,0}(x)) \sigma(t) dt,$$

$$M_{j_{E \cup \{\mu_1\}}} f(x) = \psi_0(x) \int f(\widehat{\gamma}_{t,2^{k\mu_1}-j^{\mu_1}}(x)) \psi_0(\widehat{\gamma}_{t,2^{k\mu_1}-j^{\mu_1}}(x)) \sigma(t) dt.$$

From here it is easy to see that $M_{j_{E \cup \{\mu_1\}}} - M_{j_E}$ can play the role of $R_1 - R_2$ in this situation.

With the above outlined modifications, the proof in [SS11b] goes through to prove Proposition 11.3. We leave the details to the interested reader. This completes the proof of Proposition 11.3 and therefore the proof of Theorem 7.2.

12 A closing remark

In this paper, we put more restrictions on the classes of kernels K we considered, as compared to [SS11b]. None of these additional restrictions were essential.

In [SS11b], the class of kernels $\mathcal{K}(N, e, a, \nu)$ was allowed to depend on another parameter μ_0 ($1 \leq \mu_0 \leq \nu$). In this paper, we have restricted to the case $\mu_0 = \nu$. All of the methods in this paper transfer seamlessly over to the case of general μ_0 ; we leave such details to the interested reader.

In [SS11b], the coordinates of the dilations e_j were allowed to be elements of $[0, \infty)$, instead of \mathbb{N} . This assumption was used in the proof of Theorem 9.2, but nowhere else. To deal with more general e_j , Theorem 9.2 can be replaced by the following proposition.

Proposition 12.1. *Suppose*

$$\mathcal{S} \subseteq \mathcal{A}_N^n \times [0, \infty)^\nu$$

is such that for every M , the set

$$\mathcal{C}_M := \{c \in [0, M] : \exists (Y, d_0) \in \mathcal{S} \text{ with some coordinate of } d_0 \text{ equal to } c\}$$

is finite. Then there exists a finite subset $\mathcal{F} \subseteq \mathcal{S}$ such that every $(g, e) \in \mathcal{S}$ can be written in the form,

$$g(x) = \sum_{\substack{(f,d) \in \mathcal{F} \\ d \leq e}} c_{(f,d)}(x) f(x); \quad (12.1)$$

where $c_{(f,d)} \in \mathcal{A}_N$, and $d \leq e$ means that the inequality holds for each coordinate. The neighborhood on which (12.1) holds may depend on (g, e) .

Proof. The proof proceeds by induction on ν . The base case, $\nu = 0$, follows directly from Proposition 9.5. We assume we have the result for $\nu - 1$ and prove it for ν .

Let \mathcal{M} be the module generated by $\{Y : \exists (Y, d) \in \mathcal{S}\}$. By Proposition 9.5, \mathcal{M} is finitely generated. Take $(X_1, d_1), \dots, (X_r, d_r) \in \mathcal{S}$ such that X_1, \dots, X_r generate \mathcal{M} . Define $M = \max_{1 \leq l \leq r} |d_l|_\infty$, and let c_1, \dots, c_L be an enumeration of \mathcal{C}_M . Define,

$$\mathcal{S}_0 := \{(Y, d) \in \mathcal{S} : \text{every coordinate of } d \text{ is } > M\},$$

and for $1 \leq \mu \leq \nu$, $1 \leq l \leq L$,

$$\mathcal{S}_\mu^l := \{(Y, d) \in \mathcal{S} : \text{the } \mu \text{ coordinate of } d \text{ equals } c_l\}.$$

By our assumption on \mathcal{S} ,

$$\mathcal{S} = \mathcal{S}_0 \cup \left[\bigcup_{\mu=1}^{\nu} \bigcup_{l=1}^L \mathcal{S}_\mu^l \right]. \quad (12.2)$$

Note that every $(Y, e) \in \mathcal{S}_0$ can be written in the form,

$$Y = \sum_{d_j \leq e} c_j X_j, \quad (12.3)$$

by our construction of \mathcal{S}_0 .

We apply our inductive hypothesis to \mathcal{S}_μ^l (which we may think of as a subset of $\mathcal{A}_N^n \times [0, \infty)^{\nu-1}$ by suppressing the μ th coordinate of d for each $(Y, d) \in \mathcal{S}_\mu^l$, since we know it to be equal to c_l). We therefore obtain a finite subset $\mathcal{F}_\mu^l \subseteq \mathcal{S}_\mu^l$, as in the conclusion of the proposition (with \mathcal{S} replaced by \mathcal{S}_μ^l).

By (12.3) and (12.2) it is immediate to verify that

$$\{(X_1, d_1), \dots, (X_l, d_l)\} \cup \left[\bigcup_{\mu=1}^{\nu} \bigcup_{l=1}^L \mathcal{F}_\mu^l \right]$$

satisfies the conclusion of the proposition. \square

References

- [Bou89] J. Bourgain, *A remark on the maximal function associated to an analytic vector field*, Analysis at Urbana, Vol. I (Urbana, IL, 1986–1987), London Math. Soc. Lecture Note Ser., vol. 137, Cambridge Univ. Press, Cambridge, 1989, pp. 111–132. MR MR1009171 (90h:42028)
- [Chr92] Michael Christ, *The strong maximal function on a nilpotent group*, Trans. Amer. Math. Soc. **331** (1992), no. 1, 1–13. MR MR1104197 (92j:42018)
- [CNSW99] Michael Christ, Alexander Nagel, Elias M. Stein, and Stephen Wainger, *Singular and maximal Radon transforms: analysis and geometry*, Ann. of Math. (2) **150** (1999), no. 2, 489–577. MR MR1726701 (2000j:42023)
- [Gal79] André Galligo, *Théorème de division et stabilité en géométrie analytique locale*, Ann. Inst. Fourier (Grenoble) **29** (1979), no. 2, vii, 107–184. MR MR539695 (81e:32009)
- [Lob70] Claude Lobry, *Contrôlabilité des systèmes non linéaires*, SIAM J. Control **8** (1970), 573–605. MR MR0271979 (42 #6860)
- [Nag66] Tadashi Nagano, *Linear differential systems with singularities and an application to transitive Lie algebras*, J. Math. Soc. Japan **18** (1966), 398–404. MR MR0199865 (33 #8005)
- [NW77] Alexander Nagel and Stephen Wainger, *L^2 boundedness of Hilbert transforms along surfaces and convolution operators homogeneous with respect to a multiple parameter group*, Amer. J. Math. **99** (1977), no. 4, 761–785. MR MR0450901 (56 #9192)
- [SS11a] Elias M. Stein and Brian Street, *Multi-parameter singular Radon transforms*, 2011, to appear in Math. Res. Lett.
- [SS11b] ———, *Multi-parameter singular Radon transforms II: the L^p theory*, 2011, preprint.
- [Ste80] P. Stefan, *Integrability of systems of vector fields*, J. London Math. Soc. (2) **21** (1980), no. 3, 544–556. MR 577729 (81h:49026)
- [Str11a] Brian Street, *Multi-parameter Carnot-Carathéodory balls and the theorem of Frobenius*, 2011, to appear in Rev. Mat. Iberoamericana.
- [Str11b] ———, *Multi-parameter singular Radon transforms I: the L^2 theory*, 2011, to appear in Journal d’Analyse Mathématique.
- [ZS75] Oscar Zariski and Pierre Samuel, *Commutative algebra. Vol. II*, Springer-Verlag, New York, 1975, Reprint of the 1960 edition, Graduate Texts in Mathematics, Vol. 29. MR MR0389876 (52 #10706)

MCS2010: Primary 42B20, Secondary 42B25, 26E05, 32B05

Keywords: Calderón-Zygmund theory, singular integrals, singular Radon transforms, real analytic surfaces, Weierstrass preparation, maximal Radon transforms, Littlewood-Paley theory, product kernels, flag kernels, Carnot-Carathéodory geometry